

L'ANALYSE FACTORIELLE DES CORRESPONDANCES ET L'ARBRE DE LONGUEUR MINIMUM; EXEMPLES D'APPLICATION

par J.-P. BRIANE, J.-J. LAZARE, G. ROUX et C. SASTRE¹

SUMMARY: Two numerical methods have been applied in examples of phytosociological and taxonomic studies. Those are the correspondance factor analysis and the minimum spanning tree used complementarily. The first method uses the metric of χ^2 whose interest lies in the simultaneous representation of objects and variables in the same diagram. The second is a classification method whose result is a tree which summarizes the relations between objects.

The relationships between the genus *Sauvagesia* (Ochnaceae) and other genera closely related taxonomically is given as an example of application in taxonomy.

In the phytosociology exemple, both methods are applied in the study of *Festucetum halleri* in Southern Alps. Problems of duplication of the units are discussed.

* * *

Bien que différant par leurs objectifs, phytosociologie et taxonomie sont non seulement scientifiquement complémentaires mais aussi très comparables par les méthodes utilisées. En particulier, toutes deux doivent affronter le même problème : la classification d'objets (relevés du phytosociologue, *taxa* du taxonomiste) en fonction de variables (listes floristiques du phytosociologue, caractères du taxonomiste).

Depuis longtemps, des techniques numériques simples (métroglyphes, analyses de CZEKANOWSKI, dendrogrammes, etc.) ont été utilisées par les phytosociologues et les taxonomistes. Ces méthodes sont parfois imprécises, voire entachées d'arbitraire.

L'emploi d'ordinateurs puissants permet aujourd'hui d'envisager l'utilisation de méthodes plus efficaces telles que l'analyse multivariable et les algorithmes de classification automatique.

En outre, l'augmentation des performances des ordinateurs permet de traiter des données de plus en plus nombreuses.

Deux des méthodes essayées nous ont donné particulièrement satis-

1. J.-P. BRIANE, J.-J. LAZARE, G. ROUX : Laboratoire de Taxonomie végétale expérimentale et numérique de l'Université de Paris-Sud (Centre d'Orsay) associé au C.N.R.S.

C. SASTRE : Laboratoire de Phanérogamie, Muséum national d'Histoire naturelle, Paris.

faction : l'analyse factorielle des correspondances (B. CORDIER, 1965) utilisée conjointement avec l'arbre de longueur minimum (J.-G. GOWER et G.-J.-S. ROSS, 1969). Nous allons décrire ces deux techniques et en donner des applications.

Les recherches ont été effectuées au Laboratoire de Taxonomie végétale expérimentale et numérique¹ (Université Paris-XI) en collaboration avec le Laboratoire de Statistique Mathématique (Université Paris-VII). L'ordinateur utilisé est l'UNIVAC 1108 du Centre de calcul de l'Université Paris-XI.

I. — DESCRIPTION DES MÉTHODES²

INTRODUCTION

Étant donné un tableau à double entrée dont chaque ligne correspond à un objet (relevé, taxon) et chaque colonne à une variable (espèce, caractère), on cherche à classer les objets en fonction d'affinités à partir des variables; ce tableau constitue la matrice de données.

Géométriquement chaque objet peut se représenter par un point d'un espace à m dimensions noté R^m (m = nombre de variables).

Pour comparer les objets entre eux, il est nécessaire de se donner une mesure des ressemblances en fonction des variables, aussi on fait choix d'une métrique euclidienne permettant de calculer les distances entre les points de R^m , et en particulier entre les différents points-objets. De façon pratique, plus la distance qui sépare deux points-objets est courte, plus les objets qu'ils représentent sont ressemblants.

L'ANALYSE FACTORIELLE

Dans R^m , l'ensemble des points-objets forme un nuage présentant des directions d'allongement privilégiées; l'analyse factorielle a pour but de déterminer ces directions. Dans un espace euclidien, la projection orthogonale d'un segment de droite sur un axe donné est plus courte que le segment lui-même (fig. 1).

Donc, si on projette le nuage de points-objets sur un axe, on obtient une représentation unidimensionnelle dont le principal effet est de réduire en général les distances entre points-objets projetés.

De plus, la disposition relative des points-objets projetés ne respecte plus les degrés de ressemblance des objets qu'ils représentent ex. fig. 2.

En conséquence, il faut trouver l'axe de projection déformant le moins possible le nuage initial; ce sera celui qui rend maximum le moment d'inertie du nuage par rapport à cet axe.

Ce critère est, à quelques variantes près, la moyenne des carrés des

1. — Nous sommes reconnaissants au Professeur M. GUINOCHEZ d'avoir su stimuler au sein de son équipe l'emploi de méthodes numériques nous permettant ainsi de réaliser ce travail.

2. — Par J.-P. BRIANE.

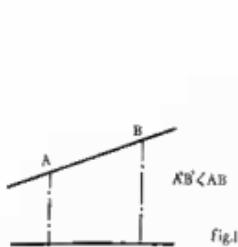


Fig.1

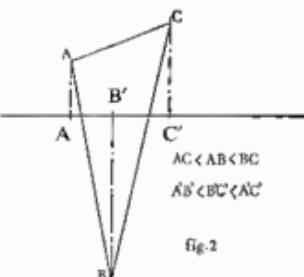


Fig.2

distances entre les points du nuage projeté. Pratiquement, plus le nombre obtenu sera grand, plus la figure de points projetés se rapprochera de la figure initiale.

En outre, le moment d'inertie I_Δ par rapport à un axe Δ est toujours inférieur au moment d'inertie total du nuage, I_{TOT} , c'est-à-dire à la moyenne des carrés des distances entre ses points.

Le moment d'inertie total I_{TOT} est une mesure commode de l'information contenue dans le tableau de données (information au sens général et non au sens probabiliste). Le taux d'inertie d'un axe $\tau_\Delta = \frac{I_\Delta}{I_{TOT}}$ représente la part d'information exprimée par cet axe.

L'axe qui rend τ_Δ maximum est appelé premier axe factoriel et son moment d'inertie, première valeur propre.

Il est très rare qu'un ensemble de données se ramène à une seule dimension. Le premier axe factoriel ne suffit pas à expliquer à lui seul les relations entre objets, il n'extract qu'une partie de l'information utile mesurée par τ_Δ . Il est donc nécessaire d'extraire plus d'information. On peut par exemple chercher un 2^e axe de projection orthogonal au premier rendant maximum le moment du nuage projeté. On calcule ainsi le 2^e axe factoriel qui extrait une part d'information plus faible que le premier.

Étant orthogonaux, les deux premiers axes factoriels extraient chacun des informations non corrélées.

On démontre qu'en règle générale on peut extraire une suite d'axes factoriels de valeurs propres décroissantes en nombre égal à celui des variables.

Les résultats se présentent sous la forme de diagrammes-plans où chaque objet est figuré par un point repéré par ses coordonnées suivant deux axes factoriels. En prenant les axes 1 et 2, on obtient ainsi la meilleure représentation plane de l'ensemble des objets.

L'ANALYSE FACTORIELLE DES CORRESPONDANCES

Jusqu'ici nous n'avons pas précisé de métrique particulière. L'originalité de l'analyse factorielle des correspondances réside dans l'emploi de la métrique du χ^2 .

Matrice des données		Poids de l'objet N° i :	
		$k_{i.} = \sum_{j=1}^m k_{ij}$	
		Poids de la variable N° j :	
		$k_{.j} = \sum_{i=1}^n k_{ij}$	
		Total des éléments de la matrice :	
		$k_{..} = \sum_{i=1}^n \sum_{j=1}^m k_{ij}$	
N° objets	$\left\{ \begin{array}{c ccc} & & \text{N° variables} & \\ & 1 & \dots & j & \dots & m \\ \hline 1 & k_{11} & \dots & k_{1j} & \dots & k_{1m} \\ \vdots & k_{21} & \dots & k_{2j} & \dots & k_{2m} \\ n & k_{n1} & \dots & k_{nj} & \dots & k_{nm} \end{array} \right.$	Carré de la distance entre les objets i et i' (suivant métrique de χ^2) :	$d^2(i, i') = k_{..} \sum_{j=1}^m \frac{1}{k_{.j}} (k_{ij}/k_{i.} - k_{i'j}/k_{i'.})^2$
Fig. 3.			

Commentons cette formule.

Tout d'abord elle ne fait pas intervenir les valeurs brutes k_{ij} de la matrice de données mais les profils $\frac{k_{ij}}{k_{i.}}$.

On élimine ainsi dans la comparaison des objets entre eux un effet de taille.

Par exemple dans le cas d'un tableau phytosociologique k_{ij} est le nombre d'espèces du relevé i . Si un relevé i possède en commun avec d'autres relevés un même nombre d'espèces, l'emploi des profils a pour effet de rapprocher i du relevé possédant le moins d'espèces.

Pour chaque variable j dans le terme correspondant de la distance intervient $k_{.j}$ en dénominateur; ceci a pour but d'éliminer un effet de taille des variables.

Par exemple pour un tableau phytosociologique où $k_{.j}$ est la fréquence absolue de l'espèce j dans l'ensemble des relevés considérés, cette opération diminue l'importance d'une espèce très fréquente donc à écologie indifférente.

Pour un tableau taxonomique où $k_{.j}$ représente la fréquence d'un caractère qualitatif j au sein de l'échantillon considéré, $k_{.j}$ très grand correspondra donc à un caractère peu discriminant. Si j est un caractère quantitatif, par exemple la mesure d'un organe, $k_{.j}$ représente alors non plus une fréquence mais un ordre de grandeur de cette mesure dont la signification n'est pas intéressante en soi. A noter qu'en analyse en composantes principales, on corrige ces inconvénients en centrant et réduisant les variables.

De plus l'usage de la métrique du χ^2 présente un avantage majeur : la représentation simultanée des objets et des variables. On peut remarquer que tout ce que nous avons dit au sujet des objets peut s'appliquer aux variables. L'ensemble des variables peut se représenter par un nuage de points d'un espace R^n (n étant cette fois le nombre d'objets); nuage dont il est possible de rechercher les axes factoriels. Cependant la symétrie est telle entre les deux nuages en usant de la métrique du χ^2 qu'il existe une relation mathématique entre les axes factoriels du nuage de points-objets et

ceux du nuage de variables. Plus précisément, à chaque axe factoriel du nuage des objets correspond un axe factoriel du nuage des variables de même valeur propre. Grâce aux formules de transition qui concrétisent ces relations, il est possible de représenter les objets et les variables sur un même diagramme. Toujours grâce à ces formules, chaque point-variable vient se placer au sein du groupe de points-objets qu'il caractérise le plus. Ceci permet une interprétation rapide et sûre de la formation de groupes ou paquets de points sur chaque diagramme.

Il est possible de chiffrer l'importance de chaque variable suivant un axe par sa contribution absolue à la valeur propre (la part d'information exprimée par l'axe, dont elle est responsable).

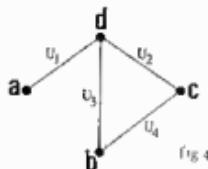
$G_k(j)$ = coordonnée de la variable j suivant l'axe k ,
contribution absolue de j à k :

$$C(j, k) = \frac{k_{..}}{k_{..}} (G_k(j))^2$$

on définit aussi la contribution relative d'une variable j comme le carré du cosinus de l'angle fait par le vecteur représentant j avec la direction de l'axe. Si cette quantité (comprise entre 0 et 1) est proche de la valeur 1, on peut dire que la variable j est très bien représentée par l'axe en question.

L'ARBRE DE LONGUEUR MINIMUM

QUELQUES DÉFINITIONS DE LA THÉORIE DES GRAPHS. On appelle « *graphe* » un ensemble de points reliés par des *arêtes* (cf. fig. 4).



Un graphe est complet si chaque paire de points est reliée par une arête (cf. fig. 5).

On appelle chemin d'un graphe, une séquence d'arêtes contiguës. Dans le graphe de la figure 4, la séquence d'arêtes u_1, u_3, u_4 constitue un chemin d'extrémités a et c . Un circuit est un chemin dont les extrémités coincident, exemple : u_2, u_4, u_3 (fig. 4). Un graphe est connexe s'il existe un chemin entre deux quelconques de ses points (cf. le graphe de la fig. 4). Le graphe de la figure 6 n'est pas connexe.

Un arbre est un graphe à la fois connexe et sans circuit (fig. 7).

Enfin un graphe valué est un graphe dont chaque arête est affectée d'un nombre (positif pour nos applications) mesurant sa longueur. Un

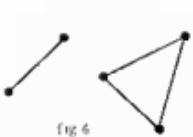


fig. 6

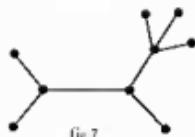


fig. 7

exemple de graphe valué est fourni par une carte routière où chaque point représente une ville et chaque arête une route.

APPLICATION A LA CLASSIFICATION AUTOMATIQUE. Dans l'espace des objets (cf. chapitre précédent) on peut représenter le nuage des points-objets par un graphe valué complet dont chaque point représente un objet et dont chaque arête a pour longueur la distance qui sépare les deux points qu'elle relie. Ce graphe comporte $\frac{n(n-1)}{2}$ arêtes (n étant le nombre d'objets).

Tenter de le dessiner est impossible. Cependant, on peut chercher à le simplifier en en retenant seulement les arêtes les plus courtes (celles qui relient des objets très semblables). Par exemple, un arbre obtenu à partir de ce graphe ne contiendra que $n-1$ arêtes. De plus cette figure possède la propriété de pouvoir s'étaler sur un plan sans que les arêtes se chevauchent.

Parmi tous les arbres qu'il est possible de construire ainsi, on choisit l'*arbre de longueur minimum*, c'est-à-dire celui dont la somme des longueurs des arêtes est minimum. On démontre que ce dernier est en général unique. Cet arbre fournit une sorte de squelette du nuage des points objets.

La construction de l'arbre de longueur minimum confiée à un programme se déroule de la façon suivante. Partant d'un point quelconque b (fig. 8) on relie ce dernier au point c qui en est le plus proche (le choix du point de départ n'importe pas). Ces deux points reliés par une arête de longueur $d(b, c)$ forment le sous-arbre A. On désigne par B l'ensemble

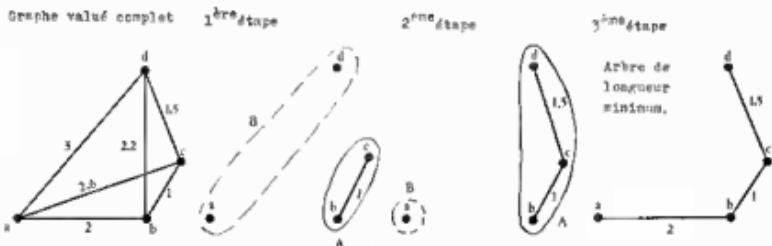


Fig. 8. — Construction progressive de l'arbre de longueur minimum.

des points non encore reliés au sous-arbre A. A chaque étape de l'algorithme on relie un nouveau point de B à un point de A de telle façon que l'arête qui les relie soit la plus courte possible. Progressivement le sous-arbre A s'accroît au détriment de l'ensemble B. La construction s'arrête

lorsque B est vide c'est-à-dire quand tous les points sont reliés à l'arbre. On peut démontrer que l'arbre ainsi construit est l'arbre de longueur minimum.

L'arbre de longueur minimum se ramifie en branches plus ou moins importantes correspondant à des groupes d'objets homogènes. Il est toujours intéressant de comparer l'arbre de longueur minimum aux diagrammes d'analyse de correspondances, il permet parfois de reconnaître un groupe d'objets homogènes qui ne se serait individualisé qu'en recourant à des axes factoriels d'ordre élevé.

Cependant dans le cas d'objets très proches les uns des autres à distances relatives sensiblement égales, la condition d'absence de circuit entraîne à des coupures arbitraires. Ainsi deux objets en réalité proches peuvent alors se trouver assez éloignés sur l'arbre de longueur minimum. Cet inconvénient n'existe pas pour l'analyse factorielle des correspondances dont les résultats plus nuancés sont donc plus proches de la réalité.

Comme la plupart des techniques de classification automatique l'arbre de longueur minimum peut s'appliquer avec n'importe quel type de distance, en particulier avec l'indice de JACCARD quand il s'agit de données phytosociologiques.

II. — EXEMPLE D'APPLICATION EN TAXONOMIE : LE GENRE *SAUVAGESIA* L. (*OCHNACEAE*)¹

L'analyse factorielle des correspondances a déjà été utilisée en systématique pour des problèmes se situant à des niveaux spécifiques ou subspécifiques (BLAISE, 1969; BRUNERYE et all., 1969; CARTIER, 1972; GARDOU, 1972).

Il nous a semblé intéressant de traiter ici d'un problème situé au niveau génétique, par exemple le genre *Sauvagesia* L.

Jusqu'à présent la systématique de ce genre était considérée comme très complexe, les relations entre les taxa infragénériques souvent mal comprises. De plus, d'autres genres très proches à limites systématiques imprécises ont été décrits : *Lavradia*, *Neckia*, *Pentaspatella*, *Leitgebia*, *Roraimanthus* et *Vausagesia*. Nous avons traité les espèces, sous-espèces et variétés, soit 35 Unités taxonomiques opérationnelles (U.T.O.) en les considérant comme faisant partie d'un ensemble taxonomique.

Pour la majorité des U.T.O., les observations portent sur 50-150 spécimens d'herbier. Malheureusement, pour certaines d'entre elles, nous avons dû nous contenter d'un nombre plus restreint de spécimens (moins de 10).

26 caractères distinctifs soit quantitatifs, soit qualitatifs, subdivisés en 80 états ont été utilisés. Les caractères quantitatifs ont été subdivisés en classes contrairement à ce qui a été fait dans certains travaux antérieurs (BLAISE, 1969; GARDOU, 1972).

Les U.T.O. ont été codées par ordre alphabétique d'espèces de 1 à 35 et les états de caractères de 1 à 80. La liste des U.T.O. et des caractères a été publiée dans un travail antérieur² (SASTRE, 1971).

1. — Par C. SASTRE.

2. — Sauf pour *S. paucielata* Sastre (U.T.O. 35), espèce décrite ultérieurement.

Les problèmes relatifs au choix des caractères ont été abordés dans ce même travail, aussi je me bornerai à résumer les raisons pour lesquelles, certains caractères n'ont pas été utilisés : ils sont redondants, ils n'ont pas été observés chez toutes les U.T.O., ils sont trop variables chez une ou plusieurs U.T.O., ils sont liés à d'autres caractères.

INTERPRÉTATION DE L'« ARBRE DE LONGUEUR MINIMUM »

Les U.T.O. se répartissent en 7 groupes (fig. 9).

Il est intéressant de noter que la majorité des U.T.O. se trouve distribuée sur l'arbre en fonction de leur adaptation morphologique et anatomique

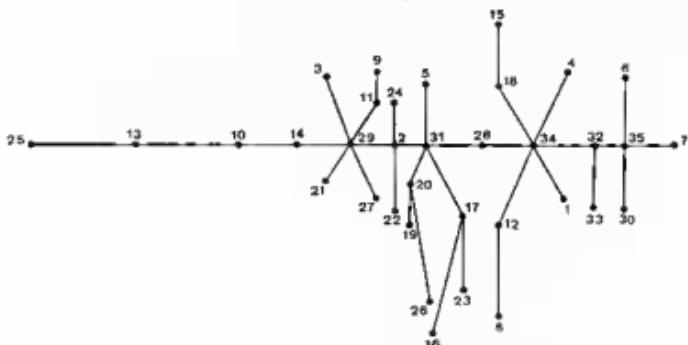


Fig. 9. — « Arbre de longueur minimum ».

à la sécheresse. Les U.T.O. les moins bien adaptées sont en majorité situées sur l'axe de l'arbre à l'exception de la partie gauche (avec 25, 13 et 10), tandis que les espèces bien adaptées se répartissent à la périphérie de l'arbre.

L'avantage de cette méthode est sa rapidité. L'« arbre minimum » prend principalement en considération les distances taxonomiques les plus faibles, or, il peut être intéressant de connaître les distances moyennes pour étudier d'une façon plus précise les rapports entre les sous-groupes.

Utilisée complémentairement, l'analyse factorielle des correspondances permet de pallier à cet inconvénient et de préciser les rapports entre les sous-groupes et quelquefois entre les sous-sous-groupes.

INTERPRÉTATION DES DIAGRAMMES D'ANALYSE DE CORRESPONDANCE

Recherchons sur le plan 1-2 (fig. 12), les groupes définis après examen de l'« arbre de longueur minimum ».

35, 6, 7, 30 se trouvent groupées avec 32 et 33; par contre dans le

plan 1-3 (fig. 10), ces deux ensembles d'U.T.O. sont séparés. Cette remarque permet de nous suggérer une relative indépendance de ces deux ensembles.



Fig. 10. — Diagramme d'analyse des correspondances des U.T.O., axes 1 et 3.

La disposition des U.T.O. dans le sous-nuage comprenant 35, 6, 7, 30 peut nous laisser sous-entendre le sens évolutif de ce groupe. L'U.T.O. 30 (genre monospécifique *Neckia*), paraissant la plus primitive est située près du point d'origine tandis que l'U.T.O. 35, plus évoluée est dans la partie périphérique du nuage.

15, 18, 34, 4, 12, 8 et 1 (genres *Lavradia* et *Vausagesia*) se retrouvent toujours dans le même sous-nuage quelque soit les axes considérés. La position des U.T.O. 1 et 34 montre leur importance comme éléments intermédiaires entre le groupe central et les autres U.T.O. incluses dans ce sous-nuage. La disposition des espèces nous suggère le sens évolutif de ce sous-groupe avec les U.T.O. 8 et 12 comme fin de phylum. Dans cet ensemble taxonomique, 8 et 12 semblent effectivement les U.T.O. les plus évoluées pour de nombreux caractères (diminution de la taille des feuilles et de l'importance de l'appareil vasculaire foliaire, passage de la placentation pariétale vers la placentation basale).

5, 31, 20, 19, 26, 17, 16 et 23. Ce groupe forme un sous-nuage, mais dans les plans 1-3 et 2-3, il existe deux sous-ensembles à direction parallèle, un premier sous-groupe composé de 5, 31, 20, 19 et 26 part toujours du point d'origine avec 5 et 31 tandis que 26 est toujours à la périphérie, un deuxième sous-groupe composé de 23, 17 et 16 avec cette dernière U.T.O. en périphérie.

Dans les différents plans, la distance entre 16 et 26 est constante.

L'étude détaillée de ces deux espèces montre qu'elles sont arrivées approximativement au même stade évolutif et qu'elles appartiennent à deux phylums différents. Elles vivent toutes les deux dans des conditions écologiques identiques. Elles doivent résister à des variations climatiques brusques : insolation, vent, dessiccation. Toutes les deux possèdent une



Fig. 11. — Diagramme d'analyse des correspondances des états de caractère, axes 1 et 2.

cuticule épaisse, des éléments lignifiés importants, et, de plus, les fleurs sont dépourvues de staminodes externes, leurs staminodes internes sont réduits et leur placentation est basale.

Il est possible de faire une remarque identique avec les U.T.O. 17 (genre *Roraimanthus*) et 23 qui se retrouvent toujours ensemble; elles appartiennent à des phylums différents et sont adaptées à des condi-

tions écologiques de savane avec brusques variations climatologiques.

22, 2, 24 avec 21, 29, 27, 3, 11, 9 et 14 forment un groupe central d'U.T.O. montrant beaucoup de ressemblances entre elles.



Fig. 12. — Diagramme d'analyse des correspondances des U.T.O., axes 1 et 2.

25, 13 et 10 ont toujours à peu près la même distance taxonomique quel que soit le plan considéré. 25 et 13 sont très proches l'une de l'autre particulièrement dans le plan 1-3 et se disposent en périphérie. Ces rapprochements sont phylogénétiquement artificiels. 25 et 13 semblent être deux fins de phylum arrivés à un niveau évolutif à peu près équivalent.

RELATIONS U.T.O. - ÉTATS DE CARACTÈRE

L'examen comparé des diagrammes d'analyse factorielle des U.T.O. et des états de caractère permet d'établir facilement les rapports existant entre les U.T.O. et ces derniers. Pour aider au raisonnement, les lignes et les colonnes de la matrice des données ont été ordonnées en tenant compte des groupements suggérés par les diagrammes d'analyse factorielle des axes 1 et 2 (fig. 12 et 11). Cette matrice remaniée sera appelée : tableau des relations U.T.O.-états de caractère.

Dans les deux diagrammes (axes 1-2 des U.T.O. et des états de caractère), le nuage obtenu adopte grossièrement la forme d'une étoile à trois branches :

— une située dans le quadrant supérieur droit du diagramme (branche 1).

	16	17	23	26	18	20	31	05	13	25	14	10	08	12	15	34	08	07	35	06	30	32	33	22	11	28	21	26	27	03	02	09	24
61	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
62	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
63	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
64	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
65	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
66	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
67	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
68	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
69	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
70	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
71	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
72	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
73	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
74	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
75	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
76	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
77	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
78	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
79	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
80	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
81	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
82	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
83	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
84	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
85	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
86	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
87	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
88	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
89	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
90	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
91	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
92	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
93	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
94	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
95	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
96	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
97	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
98	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
99	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
100	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	

Tableau des relations U.T.O.-états de caractère.

- une située dans le quadrant inférieur droit du diagramme (branche 2),
- une située dans le quadrant supérieur gauche du diagramme (branche 3).

EXAMEN DE LA BRANCHE 1

Dans la partie la plus périphérique, nous trouvons l'U.T.O. 16 et les états 61, 41 et 79.

Sur le tableau des relations U.T.O.-états, nous voyons que 61 n'existe que chez 16, 41 chez 16 et 26, 79 chez 16 et 17. Autrement dit ces trois états ne se retrouvent que chez les U.T.O. les plus périphériques et en conséquence sont bien caractéristiques de ces U.T.O.

Les états suivants (situés plus près du centre) 4, 60, 26, 85, 67... existent au moins chez l'une de ces trois U.T.O. mais aussi dans d'autres appartenant à des sous-nuages différents (à l'exception de 60 qui ne se trouve que chez 26).

Ces premières remarques sont taxonomiquement importantes. 16 est l'espèce-type du genre monospécifique *Leitgebia* qui ne possèderait comme caractère distinctif que celui de l'état 61, c'est-à-dire : épaisseur de la cuticule supérieure du limbe = 7,5-9 μ . Est-il possible de pouvoir distinguer un genre d'un autre en se basant sur ce seul critère?

Pour OLIVER (1877) ce genre devait être constitué par les U.T.O. 16 et 17. Le seul état se trouvant uniquement dans ce « genre » est le 79 : plus de 8 files cellulaires dans le massif inférieur de fibres.

GLEASON (1931) décrivit le genre monospécifique *Pentaspatella* à partir de l'U.T.O. 26. Ce « genre » ne se distinguerait du genre *Sauvagesia* uniquement par l'état 60 : épaisseur de la cuticule supérieure du limbe = 6-6,5 μ .

Les genres *Leitgebia* et *Pentaspatella* avaient été établis principalement à cause de la structure florale : placentation basale (41); absence de staminodes externes (32) et staminodes internes réduits (35).

Effectivement 41 n'existe que chez 16 et 26, par contre de nombreuses espèces de *Sauvagesia* possèdent les états de caractère 32 ou 35 et parfois 32 et 35, ex. U.T.O. 11. Pour ces raisons, il ne nous semble pas possible de maintenir ces genres.

EXAMEN DE LA BRANCHE 2

Nous trouvons les U.T.O. 8, 12, 4, 15, 18, 34 et 1, c'est-à-dire les genres *Lavradia* et *Vausagesia*; aucun état ne caractérise l'un d'eux. *Lavradia* avait été défini grâce à l'état 64 : staminodes internes soudés, or, ce caractère se retrouve toujours chez 30 (genre monospécifique *Neckia*) et souvent chez 1 (genre monospécifique *Vausagesia*).

EXAMEN DE LA BRANCHE 3

Elle est moins bien définie et est constituée principalement par les U.T.O. 7, 35, 32, 6 et 30 avec 22, 27 et 21. Aucun état de caractère n'est spécifique de ce groupe d'U.T.O.

La distribution des états dans ce tableau ne laisse apparaître aucun groupe d'entre eux susceptible de caractériser spécifiquement une ou plu-

sieurs U.T.O. Ceci nous laisse suggérer que l'ensemble taxonomique est homogène et que toutes les U.T.O. doivent être regroupées dans un genre unique, soit le genre *Sauvagesia*.

Les méthodes numériques utilisées calculent les distances taxonomiques existant entre les différentes U.T.O. Les rapprochements obtenus peuvent avoir deux origines principales : les espèces sont effectivement affines, les espèces malgré qu'elles appartiennent à des phylums différents montrent beaucoup de ressemblances, celles-ci pouvant être dues, par exemple à des convergences de forme. C'est au taxonomiste qu'il appartient de faire la distinction lui permettant ainsi de pouvoir tirer des conclusions évolutives (SASTRE, 1971).

III. — UN EXEMPLE D'APPLICATION A LA PHYTOSOCIOLOGIE : LE PROBLÈME DU *FESTUCETUM HALLERI* DU BASSIN SUPÉRIEUR DE LA TINÉE (ALPES-MARITIMES)

L'exemple utilisé nous sera fourni par le travail exécuté sur le terrain par l'un d'entre nous dans les zones de contact entre substrats calcaires et siliceux, substrats sur lesquels se différencient deux sous-associations du *Festucetum halleri* B.B. : la sous-association à *Festuca ovina* subsp. *laevis* var. *scardica* sur silice, la sous-association à *Carex sempervirens* sur calcaire (M. GUINOCHE, 1938). Les résultats de cette étude suggérée par le Professeur Guinochet à la lumière de précédents résultats (M. GUINOCHE, 1973, p. 51 et 98) ont fait l'objet d'un mémoire dans lequel l'analyse des données écologiques complète et confirme l'analyse des données floristiques résumée ici (J.-J. LAZARE, 1972).

UNITÉS OPÉRATIONNELLES ET CARACTÈRES

La méthode phytosociologique utilisée est la méthode de l'École züricho-montpelliéraise (J. BRAUN-BLANQUET, 1928, 1969). En chaque station choisie, l'observateur recense la totalité des espèces végétales présentes ou bien, comme c'est le cas dans notre exemple, les seules Phanérogames et Cryptogames vasculaires, en affectant chaque taxon (nous utiliserons par la suite indifféremment espèce ou taxon, étant bien entendu que les infraspecifiques peuvent faire l'objet de recensement séparé) d'un coefficient d'abondance. A partir de la liste floristique de l'ensemble des stations on constitue un tableau brut de données dont chaque ligne correspond à une espèce et chaque colonne à un relevé. Nous ne reviendrons pas ici sur les raisons qui nous conduisent à ne tenir compte dans les analyses que de la seule présence ou absence des espèces et non de leurs coefficients d'abondance (G. & M. ROUX, 1967; L. ORLOCI, 1968; M. GUINOCHE, 1968). Le tableau brut se présente donc sous la forme d'une matrice logique (c'est-

à-dire dont les éléments sont des 0 — absence — ou des 1 — présence — qui, signalons-le dès à présent, comporte bien plus de 0 que de 1. La forme des données se révèle donc assez identique à celle de l'exemple précédent puisque dans les deux cas, les données sont ramenées à une matrice logique.

ANALYSE FACTORIELLE DES CORRESPONDANCES

A. « ANTI-ESPÈCE » ET « ANTI-RELEVÉ »

Si une telle matrice semble se prêter particulièrement à l'analyse statistique de par son exhaustivité (pas de données manquantes) et son homogénéité, elle présente toutefois un tel « vide » (80 % de 0 dans notre exemple) que nous pourrions craindre un certain déséquilibre. Il est possible de compenser ce vide de deux manières :

— soit créer pour chaque relevé une sorte d'« anti-relevé » constitué des espèces absentes du relevé considéré, ce qui revient à associer à chaque colonne une colonne complémentaire,

— soit associer à chaque espèce une « anti-espèce » présente dans tous les relevés où l'espèce considérée est absente, ce qui adjoint à chaque ligne une ligne complémentaire.

Quels en seront les effets respectifs : dans le premier cas, la duplication conduit à donner le même poids aux espèces, dans le second cas elle donne le même poids aux relevés. Deux remarques s'imposent : d'une part l'hétérogénéité de poids des espèces et celle des relevés ne sont pas du même ordre de grandeur (pour le cas envisagé, les relevés possèdent de 14 à 41 espèces tandis que les espèces apparaissent dans 1 à 36 relevés) d'autre part « anti-espèce » et « anti-relevé » n'ont pas la même signification; en effet si l'« anti-espèce » représente la possibilité d'une espèce hypothétique de vivre en tous lieux étudiés sauf celui de l'espèce initiale (possibilité non négligeable dans un territoire homogène comme c'est le cas ici), l'« anti-relevé » représente une aberration complète de la nature puisqu'il comporterait un nombre d'espèce excessif et que la probabilité de réunir exhaustivement ces espèces est pour ainsi dire nulle.

Faut-il intervenir sur les poids? En fait poids des lignes et poids des colonnes ont des significations différentes. Une colonne est une entité constante puisqu'il s'agit d'un relevé dont le nombre d'espèces est bien défini. Une ligne, par contre, est le résultat d'une agglomération momentanée d'un certain nombre de relevés que l'on pourrait augmenter ou diminuer à loisir. Il ne paraît donc pas souhaitable de vouloir *a priori* transformer les relevés pour leur donner le même poids car ce poids a sans aucun doute une signification. Il semble par contre concevable de compenser la grande variabilité du nombre de relevés où se trouve représentée une espèce par une duplication des colonnes.

Voyons à présent les résultats de ces opérations sur l'exemple que nous avons choisi.

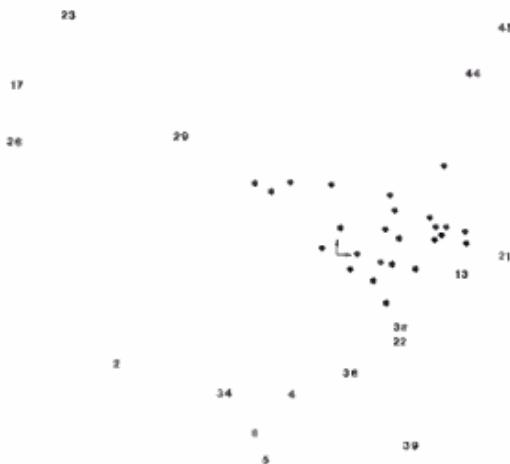


Fig. 13. — Carte des relevés à partir des données brutes; plan des axes 1 et 2 (seuls les numéros des relevés périphériques ont été précisés, les autres relevés étant figurés par un point noir pour rendre compte de leur dispersion).

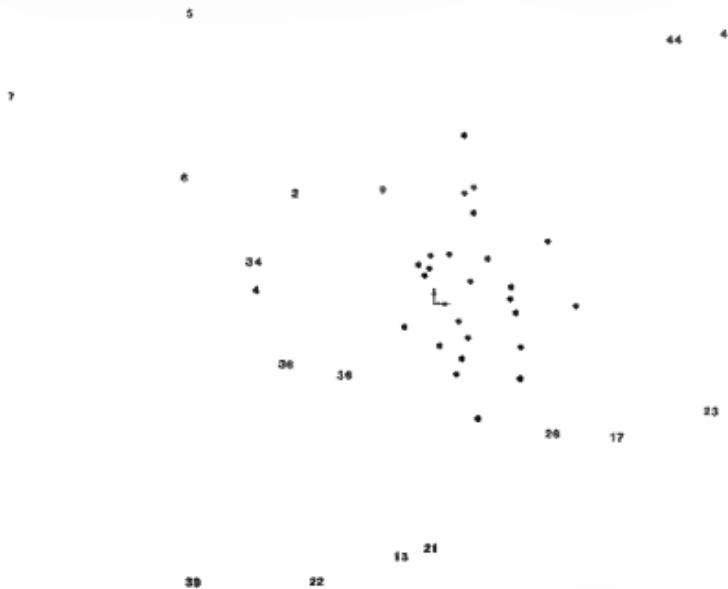


Fig. 14. — Carte des relevés à partir des données brutes; plan des axes 2 et 3.

B. GÉNÉRALITÉS SUR LES DONNÉES UTILISÉES

Les différentes analyses ont été opérées sur un tableau brut comportant 101 espèces pour 43 relevés. La numérotation de 2 à 45 provient du fait que nous n'avons pas utilisé les relevés 1 et 3 étrangers au *Festucetum halleri*. Nous insisterons sur la difficulté de circonscrire des nuages de points bien individualisés qui pourrait conduire à appuyer la théorie du « continuum » (R.-P. MC INTOSH, 1967). Cet effet est dû à l'échantillonnage délibéré de la plupart de nos stations dans des zones de contact géologique où les apports colluviaux sont importants. Les coupures se révèleraient plus aisées

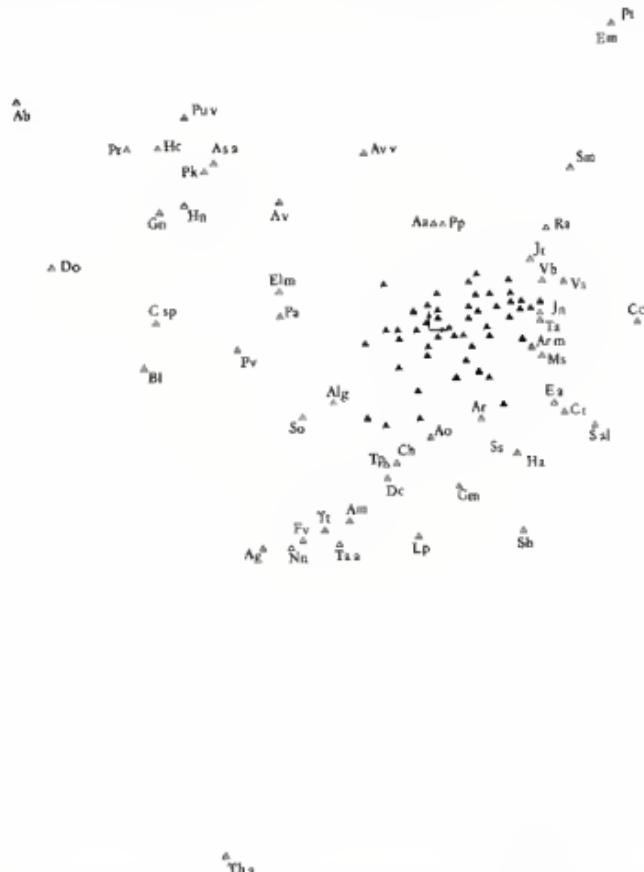


Fig. 15. — Carte des espèces à partir des données brutes; plan des axes 1 et 2 (seules les abréviations des noms des espèces périphériques ont été précisées, les autres espèces étant figurées par un triangle noir).

lors d'analyses de groupements variés d'un territoire donné, coupant ainsi court à une telle interprétation (G. ROUX, 1967, 1971; J. RITTER, 1969; A. LACOSTE, 1971, 1972; J.-C. KLEIN, 1972; M. GUINOCHE, 1973...).

C. ANALYSE DES DONNÉES BRUTES

La première analyse exécutée donne de bons éléments de réflexion. Dès les axes 1 et 2 (fig. 13) quatre directions se manifestent, matérialisées à leurs extrémités par les relevés les plus typiques, soient, compte tenu du graphe des espèces :

— 17, 23, 28 : tendance au *Seslerion* (*Festucetum halleri* sous-ass. à *Carex sempervirens* p.p.),



Fig. 16. — Carte des espèces à partir des données brutes: plan des axes 2 et 3.

- 2, 4, 5, 6, 7, 34 : tendance au *Festuceto-Trifolietum thalii* (*Festucetum halleri* sous-ass. à *Carex sempervirens* p.p.),
- 44, 45 : *Festucetum halleri* sous-association à *Festuca ovina* ssp. *lavis* var. *scardica*,
- 13, 21, 22, 36, 38, 39 : tendance au *Nardion*.

Le reste se trouve très centré autour de l'origine ce qui conduit à recourir pour la discrimination à deux méthodes : soit observer les distinctions selon les axes 3, 4 et 5, soit recourir à des analyses partielles sur les données, réduites des relevés périphériques (G. ROUX, 1971; G. & M. ROUX, 1971; C. DESARMENIEN, 1971; A. LACOSTE, 1971, 1972). Nous nous limiterons aux axes 2 et 3 (fig. 14) selon lesquels on isole facilement les relevés 13, 21, 22, 39 alors que les relevés 36 et 38 ont tendance à rejoindre le groupe 2, 4, 5... La comparaison avec les nuages espèces (fig. 15 et 16) permet d'une part de rattacher les différents ensembles à des groupements connus, d'autre part à interpréter les axes successifs d'allongement du nuage. Alors que les relevés 13, 21, 22, 39 peuvent être interprétés comme un faciès chionophile du *Festucetum halleri* avec de nombreuses espèces du *Nardion*, les relevés mal isolés correspondent à un groupement de transition entre les deux sous-associations étudiées. Compte tenu d'analyses parallèles sur des données écologiques selon la méthode précédemment décrite par A. LACOSTE & M. ROUX (1972), les axes peuvent être interprétés ainsi : axe 1 = acidité croissante, axe 2 = sécheresse et nitrophilie (corrélées ici compte tenu du mode de pâturage), axe 3 = durée d'enneigement décroissante. Il est à noter deux phénomènes : d'une part la grande contribution d'espèces de basses fréquences (présentes dans 1 ou 2 relevés), d'autre part la condensation de la plupart des relevés autour de l'origine, ce qui conduit à une discrimination difficile.

D. ANALYSE DES DONNÉES APRÈS DUPLICATION DES RELEVÉS

On remarque que sur le premier axe (fig. 17), le classement des espèces s'opère selon les fréquences croissantes. Parallèlement tous les relevés s'opposent fondamentalement au groupe des « anti relevés », ce qui se concevait de par l'anomalie qu'ils représentent. En conséquence aucune discrimination sur les critères de composition floristique ne peut se faire selon cet axe, ce qui conduit à n'utiliser pour la reconnaissance d'ensembles que les axes ultérieurs en particulier les axes 2 et 3 (fig. 18). Deux remarques s'imposent à l'observation du graphe des relevés selon ces axes : l'éclatement du nuage central dû à la condensation des « anti-relevés » autour de l'origine et la grande analogie avec l'analyse précédente puisque, mise à part la symétrie des graphes, les ensembles dégagés sont à peu près identiques. Le graphe espèce, lui, s'avère particulièrement intéressant (fig. 19) : les contributions les plus fortes se rapportent à des espèces de moyennes fréquences (6 à 36), ce qui exclut la discrimination de relevés sur la base d'espèces exceptionnelles (qu'elles soient des caractéristiques rares d'association ou des « accidentelles »). La résultante est la reconnaissance de grands ensembles de stations au détriment des très petits groupes qui s'isolent rarement. On peut pratiquer

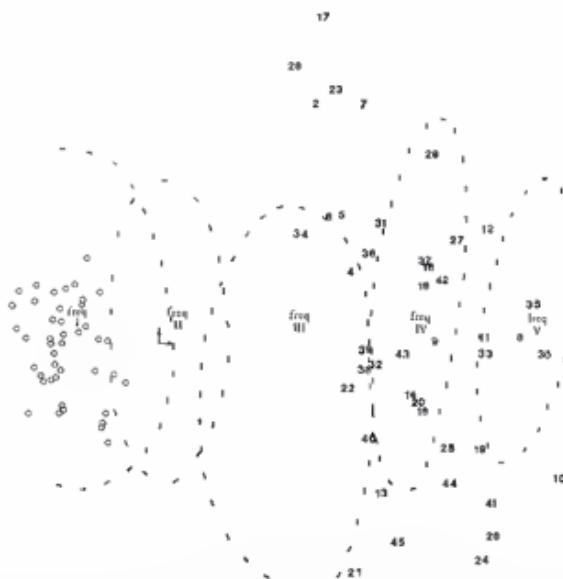


Fig. 17. — Carte des relevés après leur duplication; plan des axes 1 et 2 (les « anti-relevés » ont été figurés par un cercle blanc; les tracés en pointillé circonscrivent les points correspondant aux espèces de classes de fréquence données).

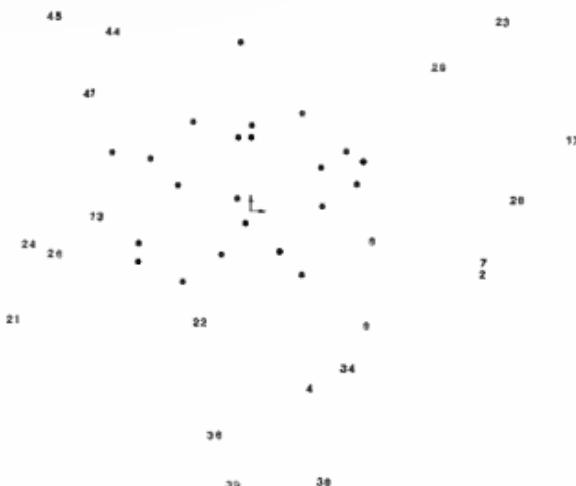


Fig. 18. — Carte des relevés après leur duplication; plan des axes 2 et 3 (les « anti-relevés » n'ont pas été figurés).

quement schématiser ces ensembles en les réduisant aux 4 quadrants, le second axe séparant les espèces des *Caricetea curvulae* à gauche, des espèces des *Elyno-Seslerietea* à droite dans leur plus grande majorité. Sur le diagramme espèce, les espèces périphériques correspondent à des caractéristiques ou des compagnes de haute fréquence pour les groupes mis en évidence :

— Axe 2 négatif, axe 3 positif = *Festucetum halleri* sous-association à *Festuca* avec notamment *Juncus trifidus*, *Sempervivum arachnoideum*, *Festuca halleri*, *Euphrasia minima*...

— Axe 2 négatif, axe 3 négatif = tendance au *Nardion* avec *Nardus stricta*, *Ranunculus montanus* et *R. pyrenaeus*, *Poa alpina*, *Geum montanum*, *Viola calcarata*...



Fig. 19. — Carte des espèces après duplication des relevés: plan des axes 2 et 3.

— Axe 2 positif, axe 3 négatif = tendance au *Festuceto-Trifolietum thalii* avec *Festuca violacea*, *Trifolium thalii*, *Trifolium pratense*...

— Axe 2 positif, axe 3 positif = tendance au *Seslerion* avec, en particulier, *Helianthemum nummularium*, *Anthyllis vulneraria*, *Gentiana verna*, *Aster alpinus*, *Elyna myosuroides*...

L'interprétation des premiers axes s'avère identique à celle de l'autre analyse, mis à part le décalage de numérotation des axes utilisés.

E. ANALYSE DES DONNÉES APRÈS DUPLICATION DES ESPÈCES

Si nous observons d'abord le nuage espèce selon les deux premiers axes (fig. 20), nous constatons la non-distinction franche espèces « anti-espèces » mais d'une manière générale une plus grande dispersion des espèces

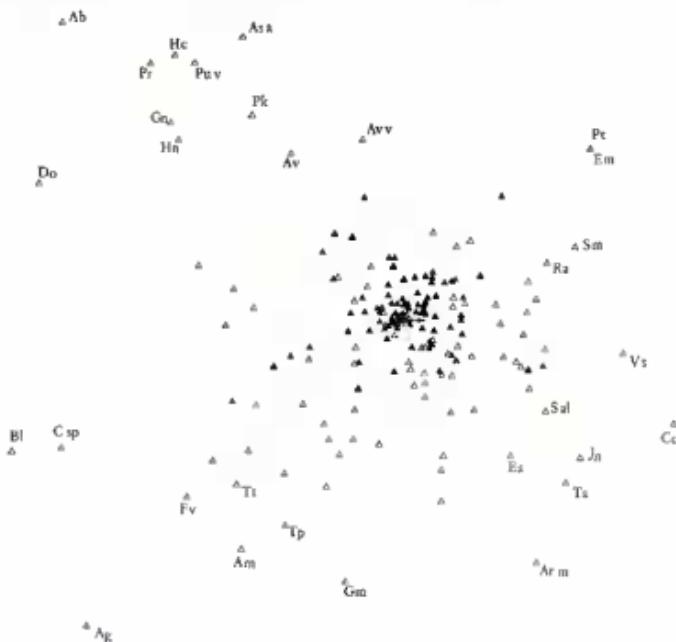


Fig. 20. — Carte des espèces après duplication de celles-ci; plan des axes 1 et 2 (les « anti-espèces » ont été figurées par un triangle noir).

que des « anti-espèces », condensées autour de l'origine. En ce qui concerne le nuage relevé (fig. 21), si la disposition est à peu près conforme à celle de la première analyse, il y a par contre accentuation de deux directions aux dépens des deux autres : l'une vers les relevés 17, 23, 28, l'autre vers les relevés 2, 4, 5, 6, 7, 34. Les plus grands changements se manifestent au niveau des relevés à grand nombre d'espèces comme le 34 (41 espèces), le 32 (40), le 4 (36), etc.

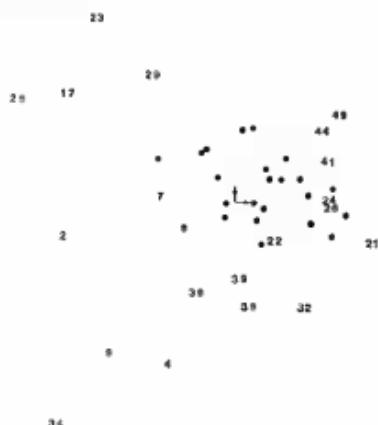


Fig. 21. — Carte des relevés après duplication des espèces; plan des axes 1 et 2.

Pour le rôle des espèces, il faut remarquer une analogie frappante avec la première analyse où les espèces de faible fréquence interviennent de manière importante, d'où les mêmes exagérations de distances pour ces

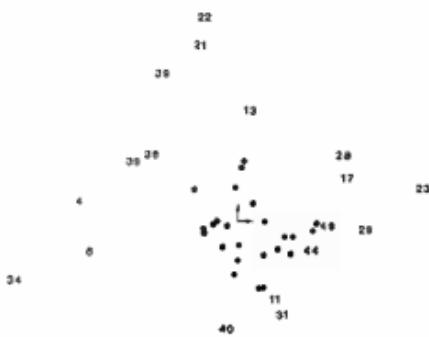


Fig. 22. — Carte des relevés après duplication des espèces; plan des axes 2 et 3.

relevés présentant des espèces rares sinon exclusives et le même tassement des autres relevés. Selon le 3^e axe (fig. 22), le groupe 13, 21, 22, 39 s'individualise comme sur la figure 14.

F. CONCLUSION SUR L'ANALYSE DES CORRESPONDANCES

Nous devons faire remarquer que les résultats obtenus dans les trois cas sont très similaires, ce qui concrétise la signification plus grande de la présence d'une espèce par rapport à son absence (G. ROUX, 1971; C. DESARMÉNIEN, 1972) et qui révèle que l'on peut se satisfaire largement, en phytosociologie, de l'analyse des données brutes. Cependant, il semble utile dans la mesure du possible, tout au moins dans le cas d'études de groupements floristiquement voisins, de faire précéder une telle analyse par une duplication des relevés pour dégager au maximum les grandes orientations du nuage en fonction des espèces de moyenne fréquence. L'analyse des données brutes permettra alors de reconnaître éventuellement des groupes de peu de relevés qui, le plus souvent, nuisent à une vision synthétique.

ARBRE DE LONGUEUR MINIMUM

Nous nous limiterons ici à quelques remarques sur le tracé obtenu (fig. 23). Nous retrouvons en fin de branches les relevés les mieux isolés par l'analyse.

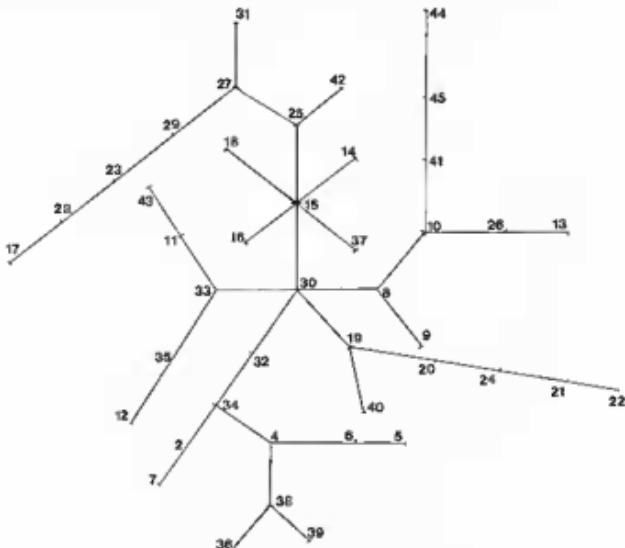


Fig. 23. — Arbre de longueur minimum tracé à partir de la distance déduite de l'indice de similarité de JACCARD.

des correspondances : (2, 4, 5, 6, 7, 34, 36, 38, 39) (17, 23, 28, 29) (21, 22), (41, 44, 45) mais, d'une part, leurs intercorrélations ne sont pas évidentes, d'autre part, les autres relevés sont liés de façon très arbitraire. Il est évident que nous nous trouvons dans le cas où la méthode est particulièrement inadaptée au type de données. En effet, les relevés utilisés ont une composition floristique très voisine et par suite, des distances au sens mathématique très proches, ce qui conduit pour obtenir une liaison unique entre relevés à trancher presque arbitrairement entre deux distances presque égales. La méthode s'opposant par définition à toute boucle s'avère ici particulièrement gênante, ce qui était bien loin d'être le cas dans les travaux phytosociologiques précédents (G. & M. ROUX, 1971; A. LACOSTE, 1972; J.-C. KLEIN, 1973; C. DESARMÉNIEN, 1971).

LÉGENDE DES GRAPHES REPRÉSENTANT LES ESPÈCES

ABRÉ- VIATIONS	ESPÈCES	ABRÉ- VIATIONS	ESPÈCES
Am	Achillea millefolium	Hc	Hippocratea comosa
Aa	Agrostis alpina	Ha	Homogyne alpina
Ar	Agrostis rupestris	Jt	Juncus trifidus
Ala	Alchemilla alpina	Jn	Juniperus nana
Af	Alchemilla flabellata	Lp	Leontodon pyrenaicus
Alg	Alopecurus gerardi	Ll	Luzula lutea
Ac	Androsace carnea	Ls	Luzula spicata
Ao	Anthoxanthum odoratum	Ms	Minuartia sedoides
Av	Anthyllis vulneraria	Mv	Minuartia verna
Arc	Arenaria ciliata	Ns	Nardus stricta
Arm	Arnica montana	Nn	Nigritella nigra
Asa	Aster alpinus	Pk	Pedicularis kerneri
Ab	Aster bellidiastrum	Pr	Pedicularis rosea
Ag	Astragalus gaudinii	Pt	Pedicularis tuberosa
Avv	Avena versicolor	Pp	Phyteuma pedemontanum
Bl	Biscutella leavigata	Ps	Plantago serpentina
Cr	Cardamine resedifolia	Po	Poa alpina
Csp	Carduus sp.	Pa	Polygala alpestris
Cc	Carex curvula	Pv	Polygonum viviparum
Ch	Chrysanthemum coronopifolium	Pc	Potentilla crantzii
Df	Deschampsia flexuosa	Pg	Potentilla grandiflora
Dn	Dianthus neglectus	Puv	Pulsatilla vernalis
Da	Draba aizoides	Rm	Ranunculus montanus
Dc	Draba carinthiaca	Rp	Ranunculus pyrenaeus
Do	Dryas octopetala	Ra	Rumex acetosella
Elm	Elyna myosuroides	Sh	Salix herbacea
Eu	Erigeron uniflorus	Sm	Saxifraga moschata
Ea	Euphrasia alpina	Sal	Sedum alpestre
Em	Euphrasia minima	Ss	Selaginella spinulosa
Fh	Festuca halleri	Sa	Sempervivum arachnoideum
Fv	Festuca violacea	Sem	Sempervivum montanum
Gb	Gentiana bavarica	Si	Senecio incanus
Gv	Gentiana verna	Se	Silene exscapa
Gm	Geum montanum	So	Soldanella alpina
Gn	Globularia nana	Stm	Statice montana
Hn	Helianthemum nummularium	Taa	Taraxacum alpinum

ABRÉ- VIATIONS	ESPÈCES	ABRÉ- VIATIONS	ESPÈCES
Tha	Thalictrum alpinum	Tt	Trifolium thalii
Ta	Thesium alpinum	Vu	Vaccinium uliginosum
Ts	Thymus serpyllum	Vb	Veronica bellidoides
Tal	Trifolium alpinum	Vs	Veronica serpyllifolia
Tp	Trifolium pratense	Vc	Viola calcarata

CONCLUSIONS

Nous espérons avoir montré le parti qu'il est possible de tirer de l'emploi judicieux des méthodes modernes d'analyse de données. Cependant, celles-ci ont des limites et elles ne sauraient remplacer le long et ardu travail du spécialiste. En effet, il convient d'apporter un soin particulier à l'étude préliminaire du problème et de définir les buts à atteindre. Ce n'est qu'après cette première étape que le chercheur commencera la collecte des données. Il est souhaitable que celle-ci s'effectue en collaboration avec un statisticien qui, en fonction des problèmes posés, conseillera une méthodologie adéquate et aidera à établir un codage efficace.

La pratique de l'utilisation des méthodes modernes d'analyse des données nous incitent à quelques remarques :

— la classification automatique, en particulier l'arbre de longueur minimum doit être appliquée avec précaution au traitement de données biologiques. L'analyse factorielle des correspondances nous enseigne que ce type de données se résume plus en gradations qu'en classes bien distinctes,

— en revanche, la classification automatique par sa simplification et sa schématisation des données peut aider à établir des coupures qui pourront éventuellement servir de base au phytosociologue et au systématicien.

BIBLIOGRAPHIE

OUVRAGES THÉORIQUES

BENZECRI J.-P. — L'analyse des données. 2 tomes, 1 248 p., Dunod. ed., Paris (1973).
CORDIER B. — L'analyse factorielle des correspondances. Thèse, 66 p. Rennes (1965).
GOWER J.-C., ROSS G.-J.-S. — Minimum spanning tree and single cluster analysis. *Applied Statistics* 18 : 54-64 (1969).
LEBART L., FÉNELON J.-P. — Statistique et informatique appliquées. 426 p., Dunod. ed., Paris (1971).
ROSS G.-J.-S. — Minimum spanning tree and printing the minimum spanning tree. *Applied statistics (Algorithms)* 18 : 103-106 (1969).

APPLICATIONS TAXONOMIQUES

BLAISE S. — Les méthodes d'analyse factorielle des correspondances appliquées à la systématique de quelques espèces de *Myosotis* L. C. R. Séances soc. Biol. 163 (1) : 83-86, 1 fig. (1969).

BRUNERYE L., GORENFLOT R. et ROUX M. — Application de l'analyse factorielle des correspondances au traitement numérique de populations de *Senecio helenitis* (L.) Cuf. Rev. gén. Bot. 76 : 217-256, fig. 1-30 (1969).

CARTIER D. — Étude biosystématique de quelques espèces du genre *Plantago* [Tourn.] L. (Section *Coronopus* DC.) II. Rev. gén. Bot. 79 : 201-248, fig. 1-20 (1972).

GARDOU C. — Recherches biosystématiques sur la Section *Jacea* Cass. et quelques sections voisines du genre *Centaurea* L. en France et dans les régions limitrophes. Feddes Repert. 83 (5-6) : 311-472, fig. 1-125 (1972).

SASTRE C. — Essai de taxonomie numérique et schéma évolutif du genre *Sauvagesia* L. *Sellowia* 23 : 9-44, fig. 1-9 (1972).

APPLICATIONS PHYTOSOCIOLOGIQUES

BRAUN-BLANQUET J. — L'École phytosociologique zuricho-montpelliéraise et la S.I.G.M.A. Comm. SIGMA n° 179 (1969).

BRAUN-BLANQUET J. & PAVILLARD J. — Vocabulaire de sociologie végétale, ed. 3, Montpellier (1928).

DÉSARMÉNIEN C. — Étude phytosociologique de relevés effectués au Chili. Mémoire ronéotypé pour l'obtention d'un D.E.A. de Statistique à l'Université de Paris, 16 p., 2 tableaux (1971).

GUINOCHEZ M. — Étude sur la végétation de l'étage alpin dans le bassin supérieur de la Tinée (Alpes-Maritimes). Thèse, Lyon, 458 p. (1938).

— Continu ou discontinu en phytosociologie. Botanical Review 34 (3) : 273-290 (1968).

— Phytosociologie. Masson ed., 227 p., 1 carte (1973).

KLEIN J.-C. — Le *Genisteto-Carlinetum macrocephalae* Ass. nova de l'étage montagnard et le *Ligusticetum corsici* Ass. nova de l'étage subalpin des massifs du Cinto et du Campitello orientale. Vegetatio 25 (5-6) : 311-333 (1972).

LACOSTE A. — La végétation de l'étage subalpin du bassin supérieur de la Tinée (Alpes-Maritimes). Thèse, Paris XI, 293 p. (1972).

LACOSTE A. & ROUX M. — L'analyse multidimensionnelle en phytosociologie et en écologie. Application à des données de l'étage subalpin des Alpes-Maritimes. I. L'analyse des données floristiques. Ecologia Plantarum 6 : 353-369 (1971).

— L'analyse multidimensionnelle en phytosociologie et en écologie. Application à des données de l'étage subalpin des Alpes-Maritimes. II. L'analyse des données écologiques et l'analyse globale. Ecologia Plantarum 7 : 125-146 (1972).

LAZARE J.-J. — Le problème du *Festucetum halleri* du bassin supérieur de la Tinée (Alpes-Maritimes). Essai d'interprétation par la méthode d'analyse factorielle des correspondances. Mémoire ronéotypé pour l'obtention d'un D.E.A. de Biologie végétale à l'Université de Paris XI, 46 p. (1972).

MAC INTOSH R.-P. — The continuum concept of vegetation. Botanical Review 33 (2) : 130-187 (1967).

ORLOCI L. — Information analysis in phytosociology: partition, classification and prediction. J. theor. Biol. 20 : 271-284 (1968).

RITTER J. — Les groupements végétaux subalpins et alpins du Vercors méridional. Thèse de 3^e Cycle, Orsay, 126 p. (1969).

ROUX G. — Un essai d'application de l'analyse factorielle à un problème de phytosociologie. Mémoire ronéotypé pour l'obtention d'un D.E.S. de Biologie végétale, Orsay, 43 p. (1967).

— Le problème des pelouses à *Nardus stricta*: une approche par les méthodes d'analyse numérique, in « Colloque interdisciplinaire sur les milieux naturels supra-forestiers des montagnes du bassin occidental de la Méditerranée » : 299-314, Édit. Centre Univ. Perpignan (1971).

ROUX G. & M. — A propos de quelques méthodes de classification en phytosociologie. Rev. Statist. Appl. 15 (2) : 59-72 (1967).

— « L'analyse des données phytosociologiques ». Communication orale, Colloque sur l'Analyse des données. Rennes (Mai 1971).